Medical Marijuana Metrics:

An Investigation into the Relationship Between the Demographic and Economic Characteristics of Oklahoman Municipalities and their Number and Concentration of Medical Marijuana Dispensaries.

Sloan C. Stinnett

May, 2020

Abstract:

The purpose of this paper was to examine how demographic characteristics of Oklahoman municipalities affect the number and concentration of licenced medical marijuana dispensaries in the municipality. To investigate this question we examined previous literature to create a host of demographic variables that we believe may be of significance to the size and concentration of the market and performed multi-linear regression to measure their effects on both. We found that in Oklahoma the main driver of the number of medical marijuana dispensaries was total population with all other demographic variables lacking significance. Poverty rates and private insurance rates were found to both have a significant and negative effect on the concentration of medical marijuana dispensaries in Oklahoman municipalities. Comparing these findings to the previous literature leads us to believe that either the dynamics between demographic variables and the size and concentration of dispensaries is sensitive to the geographic scale under study or that there is spatial non-stationarity in the relationships found for Oklahoma differ significantly from those expected when looking at studies performed on other parts of the country such as California and Colorado.

Introduction

Medical marijuana has gained ever greater support both in terms of public opinion and in state level law since the passage of Proposition 215 in California in 1996. Over the past 24 years marijuana has come to be ever more accepted as a legitimate medical good with 32 states now allowing the use of marijuana for medical purposes. One of the central businesses in the medical marijuana industry is the dispensary. Medical marijuana dispensaries serve an important function for the medical marijuana market acting as a physical location for medical marijuana patients to fill their prescriptions and purchase whatever supplies they need to consume their marijuana in the manner prescribed or that they find most effective. Often medical marijuana dispensaries also offer an assortment of marijuana strains that have differing qualities such as THC concentration and flavor profile and an assortment of novelty products related to marijuana consumption. This places medical marijuana dispensaries at an interesting crossroads between a consumer goods store and a medical supply store. Because of this unique position it becomes important for local and state governments to be aware of how demographic and economic factors affect the number and concentration of medical marijuana dispensaries so that they can effectively respond to the growing medical marijuana market and distribute resources to minimize any possible negative externalities that may come from the presence of dispensaries while at the same time insuring a minimum of access for those that rely on medical marijuana dispensaries as a source of a necessary medical good. It is with this in mind that this paper aims to test how certain key demographic and economic characteristics of a municipality relate to the number and concentration of medical marijuana dispensaries within its borders in the state of Oklahoma.

Literature Review

In order to better understand the relationship between medical marijuana dispensaries and the communities that surround them I focused on three areas of research. First, which ailments marijuana has been found to be the most useful in treating and which ailments consumers have reported using medical marijuana to treat. Second, the accessibility of pharmacies to different demographics as their similarity to medical marijuana dispensaries may also be reflected in how access is distributed. Finally, the demographic characteristics of medical marijuana consumers in states outside of our area of interest to see if they are reflected at the municipal level in Oklahoma.

Medical Marijuana: A Medical Perspective

One of the key qualities of marijuana that has been recognized by the medical community to be of great use is its ability to bring an individual into a state of hyperphagia, in which they are more willing to eat and drink (Farrimond et al, 7). This quality of marijuana has made it useful in the treatment of many eating disorders and for treating the side effects of other medicines that can oftentimes make eating unpleasant such as when chemotherapy patients are made too nauseous to eat without some form of appetite stimulant. Medical marijuana consumers have also reported that marijuana can help them control symptoms such as chronic pain, muscle spasms and anxiety (Reinarman et al, 5). However, some argue that the use cases for medical marijuana are often too broad and that this has created a large gap between the well studied uses of medical marijuana and the uses that patients have come to rely on it for (Fitzcharles et al, 5)

By working to have a comprehensive understanding of both what ailments marijuana has been shown to be the effective at treating, as well as which ailments consumers have reported as having been relieved by the consumption of marijuana, states and municipalities can take into account the medical needs of their own communities when making decisions on issues such as zoning and local cooperation with federal law enforcement on issues relating to medical marijuana.

Access to Medical Establishments

The issue of accessibility is one of greatest importance because medical marijuana dispensaries function as a centralized location for individuals to obtain medicine that in some cases is necessary for them to function normally. However, we could not find any studies that looked directly at the question of accessibility for medical marijuana dispensaries either in the state of Oklahoma or anywhere else. Because of this we decided to look to the accessibility of traditional pharmacies as they would seem to serve an analogous role for more traditional medical goods.

One study on this question was performed in East Baton Rouge, Louisiana by Ikram et al. in 2015 (Ikram et al, 1). This study measured accessibility in two distinct ways. The first measure of accessibility was travel time. The second measure was a two step floating catchment area which served the purpose of measuring how many people each pharmacy would have to serve. On average, African Americans fared better when it came to travel time but were served by pharmacies that were more responsible for serving larger populations. The opposite was true of the white population that generally had a longer travel time but were served by less crowded pharmacies. The elderly, however, fared the best with both lower average travel times and less crowded pharmacies (Ikram et al, 13-6).

This encourages us to look to see if there is a strong connection between the age and racial composition of a municipality and the number and concentration of medical marijuana dispensaries in that municipality.

Medical Marijuana Businesses and their Communities

There are three major areas of concern when it comes to the negative effects that the presence of medical marijuana dispensaries may have on local communities. First, how does the local real estate market react to the arrival of medical marijuana dispensaries. Second, are medical marijuana dispensaries associated with an increase in crime. Finally, does the presence of medical marijuana dispensaries lead children in the area to abuse marijuana.

Real Estate

Local real estate markets are very important to local communities as homes often act as the largest investments and stores of value for voters. This means that municipal and state governments have to carefully consider how changes in local law and zoning regulations will affect real estate values. For some this is a reason to restrict the entrance of medical marijuana dispensaries into their community as some view them as unsavory land uses and which will diminish the value of surrounding properties. However, the academic evidence is not clear cut on the effect that medical marijuana dispensaries have on local housing prices. Evidence from Denver, Colorado seemed to indicate that housing near dispensaries and became full retail stores actually increased in value stating that “single-family residences close to a retail conversion (within 0.1 miles) increased in value by approximately 8.4% relative to houses that are located slightly farther from a conversion (between 0.1 and 0.25 miles)" (Conklin et al, 3). However, this result does not seem to be both geographically and temporally stationary as a similar study done in Vancouver, Canada found that, "With a treatment bandwidth of 100 meters... a local marijuana dispensary reduces a home's value by an average of 3.7%. For the median valued home within the treatment radius ($490,481), the effect amounts to a reduction in home value of $18,266" (Tyndall, 15). These results are not completely comparable as one measured the effect of a retail conversion while the other simply measured based on the presence of a dispensary. However, in light of this evidence municipalities must be careful to discern how best to zone for medical marijuana dispensaries so that they minimize possible negative impact to home values while also maximizing accessibility.

One redeeming quality of the emergence of the medical marijuana market in terms of its effect on real estate values is how it benefits local warehousing rental prices. As a New York Times article from April 2017 shows that restriction on interstate transport and the strong demand for a high volume and variety of product has led many growers to search for warehouse space to use as growing facilities. The renters also generally pay higher than market value rents to compensate for the risk that comes to the land owners for renting out their property for work that exists in a legal grey zone (Gelles). With sufficient local demand these kinds of growing operations can come to compose a large part of the total warehousing market "representing roughly 5.0% of the 231 million square feet of industrial space in Metro Denver" and "1.5% of the total 74.58 million square feet of industrial space in the Boulder-Broomfield corridor" (Zhang et al. , 11-2).

These two factors should be weighed against each other for each municipality as they adjust to this emerging market. Municipalities with large amounts of warehousing and generally low rents may be incentivised to market themselves as welcoming to medical marijuana products while more suburban and family focused municipalities may want to carefully consider how best to zone their commercial areas for medical marijuana dispensaries.

Crime

Another concern for the growing medical marijuana industry is its association with criminality. As marijuana is a schedule one drug and has been one of the central targets of the war on drugs marijuana for many has an immediate association with criminality. This fact compounded with the fact that many dispensaries have to work in cash to avoid possible money laundering charges because of the federal prohibition on the sale of marijuana leads many to be concerned as to whether the presence of a medical marijuana dispensary invites more crime into the area. This again is a contentious issue in the literature. One study performed on Long Beach, California found that “Across local and adjacent areas, an increase of one dispensary per square mile was related to a 1.5–4.8% increase in violent crime” and “ a 0.4–2.6% increase in property crime” (Freisthler et al. , 5-6). One caveat to this, however, is the fact that the majority of the effect did not come from the local density of medical marijuana dispensaries but from the density of medical marijuana dispensaries in neighboring block groups (Freisthler et al. , 5-6). A similar study on South Los Angeles, California that measures effects at the census tract level found that dispensaries were not associated with increases in violent or property crime (Subica et al. 1).This conclusion is in agreement with findings from Washington D.C that found either a decreased or constant level of violent and property crime at the street level after the opening of medical marijuana dispensaries(Stephen et al, 1).

Abuse by Minors

Another area of concern for some areas is the possibility that the presence of medical marijuana dispensaries will encourage the use and abuse of marijuana by minors. However, the academic literature seems to run counter to indicate that this should not be an area of concern. One study performed on 8th through 12th graders in California found that there was no linkage between the presence or number of dispensaries, the price they sold marijuana at, or the variety of products sold at the dispensaries and adolescent consumption of marijuana in the past month or their propensity to consume it in the future (Shi et al. 2018, 3-4). Similarly, a 2016 study found that minors composed a very small percentage of all medical marijuana patients representing less than 1 percent of the total population; however, the percentage varied by state (Fairman, 1). This point to use and abuse by minors being a relative non-issue for areas that are planning for the growth of the medical marijuana industry.

What are the Demographics of Medical Marijuana Dispensary Consumers ?

The final question to be asked about this emerging medical marijuana industry is who exactly are its core consumers? Are there any particular demographic features of individuals that become medical marijuana patients or frequent medical marijuana dispensaries? A 2011 study on nine medical marijuana assessment clinics in California found that for the 1746 patients included in the study “the MM patients are three-fourths male and three-fifths White. Compared to the US Census of California, the patients in this sample are on average somewhat younger, report slightly more years of formal education, and are more often employed” (Reinarman et al.

3). However, Reinarman et al. does warn that there may be significant systematic bias in the sample selection process of this experiment. This bias comes mainly from two sources: the expense that is associated with making an appointment with the medical marijuana assessment clinic and legal concerns unique to California that may prevent Latinos from being properly represented (Reinarman et al, 3). The high concentration of male consumers is supported by more recent evidence on trends in registered medical marijuana patients that looked at data from 6 states for the years 2012 through 2015 and found that roughly two-thirds of all registered patients were male (Fairman, 4).

At the neighborhood level a study done on Los Angeles, California found that predominantly hispanic neighborhoods had a higher density of medical marijuana dispensaries. Similarly neighborhoods that had high levels of connectedness to transportation and other business, such as those that had a highway ramp or high percentage of the neighborhood zoned for commercial use also had higher densities of medical marijuana dispensaries (Thomas and Freisthler, 3). This study is corroborated by evidence out of Colorado that found a "greater availability of medical marijuana stores in neighborhoods that had a higher proportion of minority, higher level of poverty, and higher density of alcohol outlets" (Shi et al. 2016 , 5).

Data

List of Dispensaries from the Oklahoma Medical Marijuana Authority

The Oklahoma Medical Marijuana Authority (OMMA) is a branch of the Oklahoma State Department of Health created for the purpose of regulating the medical marijuana marketplace at all levels in the state of Oklahoma. OMMA is responsible for the licensing of doctors, patients, dispensaries, growers, transporters, processors, waste disposal managers and laboratories. Most importantly for this research they publish and regularly update a list of all licensed dispensaries in the state of Oklahoma

Contents

The version used for the analysis in this paper contained five pieces of information for each currently operating licensed dispensary in the state of Oklahoma as of November 20, 2019: license number, name, city, zip code and county. For the purposes of this study the information on the city that the dispensary was located in was the most important. In recent releases OMMA has added information such as trade name, email and phone number to the list of reported variables.

Usage

For the sake of this study the variable of city name was most important. By calculating the number of occurance of each city name in the list we were able to create a table consisting of the name of the city as written in the OMMA's list of licensed dispensaries and the number of licensed dispensaries in each of those cities.

Limitations

The primary limitation of this data is the fact that it only contains licensed dispensaries. It is possible that there are illegally run dispensaries that operate within Oklahoma municipalities and that they draw away business from licensed business. This means that the data cannot give us a full picture of the marijuana market place in each municipality. The data can only give us information on the legal part of the marijuana market. How large a segment of total demand for marijuana, even for medical uses, is supplied by the black market may vary between municipalities and be influenced by the factors we are interested in and, as such, may bias our results.

Preliminary Observations

We can make some preliminary observations about medical marijuana dispensaries in Oklahoma from this data alone. First, by dividing the total number of dispensaries 2169 by the number of municipalities in the state we can find the average number of dispensaries per municipality which is 3.6515. We can also see which municipalities have the most dispensaries: Oklahoma City has 501, Tulsa has 263, Edmond has 88 and Norman has 66.

Census 5-year ACS data 2018 and 1-year ACS data 2017

In order to measure demographic variables for the municipalities I turned to the 5-year American Community Survey (ACS) estimates for the year 2018 as well as the 1-year ACS estimates for the year 2017 in order to obtain estimates for median household income for some municipalities that had their estimates suppressed in the 2018 data. This data source allowed me to get information on a wide range of demographic variables for all municipalities in Oklahoma that were also in the list of incorporated places used by the census bureau. I gathered the data by making a series of calls to the census API using the census API package for R (Hannah , 2019).

Contents

The data pulled from the Census's 5-year estimates included: Total population, Median Household Income, Population White, Population in Poverty, Population 20-24, Population 25-34, Population in the workforce, Working age population, population 18-64 with private insurance and no disability, population 18-64 with private insurance and a disability, population 18-64 with public insurance and no disability, population 18-64 with public insurance and a disability. These would be refined into the final variables for the regression. Only the estimate for median household income was pulled from the 1-year ACS data for the year 2017 and only for those locations where the data was both unavailable in the 2018 data and was available in the 2017 data.

Usage

Variables such as total population and median household income were used as they appeared in the original data. However, some of the other variables where used to create calculated variables such as workforce participation, percentage of total population that does not identify as white, percent of the population that is in poverty, percent of the population that is between the ages of 20 and 34, percent of the population that has private insurance, percent of the public that has public insurance, and combined with the number of medical marijuana dispensaries found in the OMMA data a count for the number of medical marijuana dispensaries for each one hundred people living in a municipality was calculated.

Limitations

There are two major limitations to the data acquired from the 5-year American Community Survey. First, the fact that they are estimates and not observations of the true values means that from the beginning there may be some bias in our results. Second, values of certain variables, namely median household income, are suppressed for the purpose of anonymity of those being observed or because of the lack of viable observations. This means that it is possible that some of the observations that were removed because they did not have an estimate for the median household income may have been missing that value in a non-random way which would bias our results to only being applicable to the community for which the criterion of exclusion did not apply.

Map of Oklahoma's Municipalities from the Oklahoma Department of Transportation

As we are using municipalities as our unit of observation it is important to consider geography and how the municipalities relate to one another in space. As such it was necessary to find a map that properly represented municipal boundaries. The map used in this analysis came from the shape files provided by the Oklahoma Department of Transportation (ArcView).

Contents

The shape file was composed of 1413 observations many of which are for unincorporated areas. After subsetting such that only those areas with a state FIPS code matching Oklahoma's , 40, we are left with 852 observations of areas that we will refer to as municipalities. However, some of these are repeats and so we remove all the non-unique observations leaving us with only 594 municipalities.

Usage

The most important use of the map of municipal boundaries is for our test for geographic stationarity. As we are using geographic data it is important to check to make sure that our regression is equally representative of the whole geographic area under study rather. In order to do this we turn to a Moran's I test on the residuals of the model. This test essentially tells us whether or not our model is equally applicable over the entire geographic area or if it would be better to to split up different geographic areas and run separate regressions for each of them.

Limitations

The central limitation comes from how this dataset pairs up with census geography. Because municipalities are not a level of census geography but are usually included in the set of incorporated places there is not a perfect one to one match between all the incorporated places that we have data on and the municipalities that have geography for. Because of this we had to limit the scope of our investigation to only those places that are included in both the list of incorporated places and the list of municipalities.

Methods

Variable Selection

Selecting Which Variables to add to the Model

When first selecting variables we examined which demographic variables had been highlighted by the previous literature as being of importance name variables concerning race, income, and age. However, as many of these variables and their importance to predicting medical marijuana consumption or pharmacy access had been based on more micro level data than the municipality, we decided to also examine some of the key characteristics of the municipality itself namely the population, labor force participation. Finally, variables concerning the insured population of the municipality were added to see if more or less insured municipalities had different outcomes.

Selecting Which Variables to Keep in the Model

The final factor for variable selection was our goal to minimize multicollinearity in the predictors. It was this goal that led us to change all variables besides total population and median household income to their percentage of the total population of the municipality rather than as pure counts so as to avoid possible multicollinearity with total population. Next, we examined the correlation matrix in Figure 1 finding that there does not appear to be a very high correlation between any of the predictors and each model. Finally, the variable inflation factor (VIF) for each variable in each model was calculated with those with a VIF of three or greater facing special scrutiny.

Model Selection

The basic model chosen for this study was the multi-linear regression. This model was chosen for its properties as the best linear unbiased estimator and its easy interetability.

Checks for Gauss-Markov Assumptions

Linearity

Linearity of the predictors was examined by looking at the residuals vs fitted plots of the three models. These plots along with the normal Q-Q plot, scale vs location plot and residual vs leverage plot for each of the models are presented in Figures 5-7. The goal of this check was to temper our interpretation of the results of the regressions based on how closely they seem to match the expected residual vs fitted plot which would generally appear as a cloud of points with a flat horizontal line through the points.

Absence of Multicollinearity

The assumption of independence of the independent variables was checked in two ways. First, a graphical correlation matrix was created, seen in Figure 1. The correlation matrix indicated that there were no strong linear relationships between the predictors. The second check was to calculate the VIF for each model. The VIF for each variable in each model can be found in the finds section alongside other information for each model.

Heteroskedasticity

To check for heteroskedasticity a Breusch-Pagan test was run on each model and where heteroskedasticity was found the models were rerun using robust standard errors. The results of the Breusch-Pagan test for each individual model can be found in the findings section and the results of the regression both with and without robust standard errors can be found for each model in Figures 2-4.

Randomness and Exogeneity

The assumptions of randomness and exogeneity are two assumptions that we take as being given. However, there is good reason to believe that these assumptions may not be met. The assumption of randomness is quite tenuous as the data reported in the ACS 5-year and ACS 1-year estimate are subject to systematic suppression for the sake of anonymity and quality control (United States). This means that there is a high chance that observations missing from the final dataset used for the regressions had observations that are missing not at random and so may be systemically skewed. The assumption of Exogeneity is also taken as a given. It is possible that some of the variables in the model are simply correlated with unmeasured variables such as average educational attainment of the populace.

Geographic Stationarity

In order to check to see if our model performs better in some geographic areas and more poorly in others we used a Moran's I test under randomisation with an alternative hypothesis that there is clustering or dispersive effects in residuals as defined in the two sided Moran's I test under randomisation from the spdep package for the R programming language (Bivand). For this test we used a nearest neighbors definition where municipalities whose centroids were within 90km of one another were defined as being neighbors. If the Moran's I test under randomisation returned a p-value less than .20 than the model was considered to be geographically non-stationary.

Findings

Model 1

The first model was a multi-linear regression with the number of licensed medical marijuana dispensaries in the municipality as the dependent variable, denoted freq in the table, and median household income (MedHosHolInc), percent population non-white (PerPopNonWhite), percent of total population in poverty (PercentPov), percent of the total population between the ages of the twenty and thirty four (Percent20-34), labor force participation (LaborPart), percent of the total population with private insurance (PercentPrivIns) as the independant variables.

Linearity

As we can see from the residuals vs fitted graph in Figure 4 there is evidence that there is nonlinearity between predictors and the variable of interest. This may bias the results away from their true values as the estimate given does not properly account for the non-linear relationship between the predictors and the number of licensed medical marijuana dispensaries in the municipality.

VIF

Totpop	MedHosHolInc	PercentPopNonWhite	PercentPov	Percent20-34	PercentPrivIns
1.059	1.161	1.474	1.228	1.219	2.146

As we can see from the preceding table none of the variables in the model have a VIF greater than 3 and as such have a generally acceptable level of correlation with one another. However, the value of 2.146 may be of concern to some and may warrant removal of percent of the population with private insurance as a predictor if suppressing multicollinearity is one's highest concern.

Breuch-Pagan test

To test for heteroskedasticity in the first model we performed a Breush-Pagan test for heteroskedasticity with the null hypothesis of constant variance. The test returned a chi-squared value of 891.7441 with 1 degree of freedom. The probability of getting this chi-squared value assuming homoskedasticity is 6.117374e-196 which is far below the 0.05 threshold and so we default to the alternative hypothesis that there exists heteroskedasticity in the model. As we could not precisely identify the source of the heteroskedasticity we choose to rerun the model with heteroskedasticity consistent standard errors as calculated by the coeftest function from the lmtest package (Achim).

Moran’s I under randomisation

The final test that we ran on the model was a Moran’s I test under randomisation to test the hypothesis that there is geographic clustering or dispersive effects in the values of residuals for regions whose centroids are within 90km of one another which could suggest geographic non-stationarity. This test was performed on the model using the moran.test function from the spdep package with the two tailed alternative hypothesis (Bivand). The test returned a Moran's I standard deviation of -0.83475 and a p-value of 0.4039 which is well above our cut off of .20 so we fail to reject the null hypothesis that there is no spatial auto-correlation or dispersion in the model and so can move forward with a single global model.

Model

The final results for the model can be seen in Figure 2. As we can see in the original OLS model both the intercept and total population variable were significant at the 0.05 level. However, after adjusting for heteroskedasticity the total population keeps is high significance with a p-value lower than 0.01; however, the constant term now is only significant at the 0.1 level. We also see that the model has a very high adjusted R-squared of .976. From this we can see that the variable of greatest importance in explaining the number of licensed medical marijuana dispensaries in municipalities that are also incorporated places in Oklahoma is the total population of the municipality.

Model 2

Model 2 keeps the same dependent variables as model 1 and all the independent variables; however, it also adds an interaction term between total population and itself. This interaction term allows for us to account for some of how the effect of total population on the number of licensed medical marijuana dispensaries changes as total population increases and decreases.

Linearity

Looking to the residual vs fitted plot in Figure 6 we see that although the line does seem to come closer to what we would expect from a linear relationship that it still does not perfectly reflect a linear relationship and so the estimated effects may be skewed as the model does not perfectly account for the non-linear relationship between the number of licensed medical marijuana dispensaries and the independent variables .

VIF

Totpop	MedHosHolInc	PercentPopNonWhite	PercentPov	'Percent20-34'
PercentPrivIns	LaborPart	I(Totpop *Totpop)		
12.111	1.162	1.478	1.239	1.258
2.181	1.824	11.670		

From the table above we see that the only variables that are overly inflated are total population and its iteractionterm with itself. This is to be expected and should not bias the results significantly as long as the both total population and and its self-interaction term are both significant.

Breuch-Pagan test

Our Breuch-Pagan test for heteroskedasticity returned a chi-squared value of 34.742 which gives us a p-value of 3.765e-09 which is highly significant. This indicates that there is heteroskedasticity in the model and that we should use heteroskedasticity consistent standard errors. This update model can be seen in the second column from the left in Figure 3.

Moran's I under randomisation

Using the same methodology as was used for model 1 we calculated a Moran's I standard deviation of -1.175 with a p-value of 0.24. This brings it close to our cut off point for significance; however, we still do not have significant evidence that there's geographic non-stationarity in the residuals of the model and so we conclude that the global model is a sufficient representation of the effects for all the regions under study.

Model

The final results for model 2 both before and after adjusting for heteroskedasticity can be found in Figure 3. As we can see from the column labeled OLS both total population and it's self interaction term are highly significant. The intercept, however, has lost all significance it had in model 1. However, as we see in the coefficient test column after adjusting for heteroskedasticity the interaction term becomes insignificant, and percent of the total population in poverty picks up some significance at the 0.1 level. Overall this new model does not change our understanding much, still pointing to total population as the main driver in the market size but hinting that there may be some effects coming from percent of the total population that is in poverty.

Model 3

Our final model replaces the total number of medical marijuana dispensaries in a municipality with the number of medical marijuana dispensaries per 100 residents of the municipality. We also remove the total population from the set of independent variables. This lets us focus more on the concentration of licensed medical marijuana dispensaries and more closely examine the variables that have a weaker effect than total population but that may still have some explanatory power

Linearity

Looking to the residual vs fitted plot in Figure 7 we see that once again the plot indicates non-linearity and so the final results may be skewed by improperly accounting for the nonlinear relationships between the dependent and independent variables.

VIF

MedHosHolInc	PercentPopNonWhite	PercentPov	‘Percent20-34‘	PercentPrivIns	LaborPart
1.160	1.464	1.227	1.211	2.137	1.795

Following the table above it does not appear that there is significant inflation of any of the variables in the model. This suggests that there is not a high level of multicollinearity between any of the predictors being used.

Breuch-Pagan test

The Breush-Pagan test for heteroskedasticity performed on the model returned a chi-squared value of 172.3825 with a p-value 2.232e-39. With this we can confidently reject the null hypothesis of homoscedasticity. In order to correct for this we recalculate the model with heteroskedasticity consistent standard deviations using the same methodology as we did for model 1 and model 2. The results of this refined model can be seen in the coefficient test column of Figure 4.

Moran’s I under randomisation

The results of the Moran’s I test under randomisation for model three are very similar to those for model 1 with a Moran’s I statistic standard deviation of -0.834 and a p-value of 0.404.

With this we fail to reject the null hypothesis that the residuals of the model are not geographically auto-correlated and do not exhibit dispersive effects.

Model

Looking to Figure 4 we can see that in the OLS regression column we can see that percent of the total population in poverty, percent of the total population that have private health care and the constant term are all significantly different from zero at the 0.01 level. We also see that the models adjusted R-squared value is 0.079 indicating that a large amount of the variation in the number of licensed medical marijuana dispensaries per 100 citizens in an Oklahoma municipality is left unexplained. After correcting for heteroskedasticity all significant variables remain significant at the 0.05 level with only the independent variable percent of the population with private insurance losing its significance at the 0.01 level. The fact that the coefficient for percent of total population that is in poverty is significant and negative in both model 2 and model 3 suggests that municipalities that have high levels of poverty are less attractive as locations for licenced medical marijuana dispensaries to establish themselves. The fact that percent of the total population that has private insurance is also significant and negatively correlated indicates that highly insured municipalities may have less use for licensed medical marijuana dispensaries.

Discussion

The Oklahoman Medical Marijuana Market

The Oklahoma medical marijuana market, when examined at the municipal level, seems to follow the general notion of supply and demand with larger municipalities commanding larger markets. However, it would appear that municipalities that are more impoverished and have higher rates of private insurance generate less demand per person causing the concentration of dispensaries to be lower. This may be caused by a confluence of income and substitution effects. As the insured population increases a wider range of substitutes for medical marijuana is opened up and so less dispensaries are needed to fill medical demand. At the same time medical marijuana may act as a price itself out of competitiveness for sales towards impoverished individuals but this income effect may either be non-linear and concentrated at lower incomes leading to the lack of significance of the median household income variable but strong significance of poverty rates. It may also be the case that the poverty rate acts on concentration indirectly with other variables external to the model bit highly correlated with poverty such as educational attainment or crime rates influencing the concentration of medical marijuan dispensaries in these municipalities. Further research into these relationships and how they change with geographic scale and place is warranted.

Implications for zoning and city planning

The main implication for city planning is that larger cities are going to be the ones that must plan for the largest increase in the number of dispensaries. Given that poverty has a countervailing effect on concentration it would be opportune for municipalities especially those with large populations and high poverty rates to research whether the number and concentration of medical marijuana dispensaries they are allowing to open for business is enough to properly serve the needs and demands of their community especially those that may be underserved and offer incentives for new entrants if they find either lacking. However, with the evidence presented by Freisthler et al and Subica et al, offering conflicting evidence on the effect that the presence of medical marijuana dispensaries will have on crime rates, it would be prudent for cities that expect large numbers of entrants into their medical marijuana markets to invest in individual studies on the effects that early entrants have on crime in their municipalities so that they may better tailor policing resources to meet any possible increased demands for service.

Relationship to Past Research

As much of the previous research looked at the medical marjuana industry at a smaller scale than the municipality we can use this as an opportunity to examine how effects change as the geographic scale changes. The first result of interest comes from comparing our results to those of Shi et al 2016. Where Shi et al found that for Coloradoan neighborhoods that had higher rates of poverty and higher proportions of minority occupants also had better access to medical marijuana dispensaries, our evidence suggests that for Oklahoman municipalities this association does not hold either for the number of medical marijuana dispensaries or for the concentration. Instead we find that there is no evidence of correlation between municipal level minority composition and the number or concentration of medical marijuana dispensaries. Similarly our evidence suggests that for Oklahoman municipalities the association with poverty rates is inverse and that municipalities that have higher rates of poverty appear to have the same number of dispensaries altogether but a high concentration as measured by citizens per dispensary.

This disconnect between our municipality level results and the neighborhood level results found in Shi et al suggests a few courses of action. First, a finer grain analysis of Oklahoman neighborhoods to see if the same relationship found in Colorado is mirrored in Oklahoma and second, for the reverse to be done in Colorado with a more macro level investigation into the effects at the municipal level. This will give a clearer picture on whether the association is sensitive to the level of geography under study or if the effect is geographically non-stationary with each state having significant differences in their municipal level medical marijuana dispensary markets.

Conclusion

In conclusion, after examining the relationship between our host of demographic variables and their relationship with the number and concentration of medical marijuana dispensaries in Oklahoman municipalities we have found, using multi-linear regression, that total population has the greatest effect on the size of the licensed medical marijuana market in a Oklahoman municipality with factors such as household income, racial composition, and age composition having no significant role. However, poverty rates and private insurance rates were both shown to have significant and negative effects on the concentration of dispensaries in a municipality. It would appear that at the municipal level the medical marijuana market obeys the simple system of supply and demand with larger municipalities controlling larger markets. However, it would appear that there are some countervailing effects related to poverty rates and insurance coverage that have a significant impact on the concentration of dispensaries in Oklahoman municipalities and may harm overall accessibility.

References

Achim Zeileis, Torsten Hothorn (2002). Diagnostic Checking in Regression Relationships. R News 2(3), 7-10. URL https://CRAN.R-project.org/doc/Rnews/

Aravind Hebbali (2018). olsrr: Tools for Building OLS Regression Models. R package version 0.5.2. https://CRAN.R-project.org/package=olsrr

ArcView SHP Files. https://www.odot.org/hqdiv/p-r-div/maps/shp-files/. Accessed 6 Febuary 2020.

Bivand, Roger S. and Wong, David W. S. (2018) Comparing implementations of global and local indicators of spatial association TEST, 27(3), 716-748. URL https://doi.org/10.1007/s11749-018-0599-x

Conklin, James, et al. "Contact High: The External Effects of Retail Marijuana Establishments on House Prices." *Real Estate Economics*, vol. n/a, no. n/a., 2017, *Wiley Online Library*, doi:10.1111/1540-6229.12220. Accessed 7 Dec. 2019.

Fairman, Brian J. "Trends in Registered Medical Marijuana Participation across 13 US States and District of Columbia." *Drug and Alcohol Dependence* 159 (2016): 72-79. Web.

Farrimond, Jonathan A., et al. "Cannabis Sativa and the Endogenous Cannabinoid System: Therapeutic Potential for Appetite Regulation." *Phytotherapy Research*, vol. 25, no. 2, 2011, pp. 170–88. *Wiley Online Library*, doi:10.1002/ptr.3375.

Fitzcharles, Mary‐Ann, Daniel J. Clauw, Peter A. Ste‐Marie, and Yoram Shir. "The Dilemma of Medical Marijuana Use by Rheumatology Patients." *Arthritis Care & Research* 66.6 (2014): 797-801. Web.

Freisthler, Bridget, et al. “A Micro-Temporal Geospatial Analysis of Medical Marijuana Dispensaries and Crime in Long Beach, California.” *Addiction*, vol. 111, no. 6, 2016, pp. 1027–35. *Wiley Online Library*, doi:10.1111/add.13301.

Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2019). dplyr: A Grammar of Data Manipulation. R

package version 0.8.3. https://CRAN.R-project.org/package=dplyr

Hannah Recht (2019). censusapi: Retrieve Data from the Census APIs. R package version 0.6.0. https://CRAN.R-project.org/package=censusapi

Hlavac, Marek (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables. R package version

5.2.1. https://CRAN.R-project.org/package=stargazer

Ikram, Samina Z., et al. “Disparities in Spatial Accessibility of Pharmacies in Baton Rouge, Louisiana.” *Geographical Review; New York*, vol. 105, no. 4, Oct. 2015, pp. 492–510.

John Fox and Sanford Weisberg (2019). An {R} Companion to Applied Regression, Third Edition. Thousand Oaks CA:

Sage. URL: https://socialsciences.mcmaster.ca/jfox/Books/Companion/

“List of Licensed Businesses.” *Oklahoma Medical Marijuana Authority*, http://omma.ok.gov/list-of-businesses. Accessed 10 Dec. 2019.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reinarman, Craig, et al. "Who Are Medical Marijuana Patients? Population Characteristics from Nine California Assessment Clinics." *Journal of Psychoactive Drugs*, vol. 43, no. 2, Apr. 2011, pp. 128–35. *DOI.org (Crossref)*, doi:10.1080/02791072.2011.587700.

Roger Bivand and Nicholas Lewin-Koh (2019). maptools: Tools for Handling Spatial Objects. R package version
0.9-5. https://CRAN.R-project.org/package=maptools

Shi et al. - 2016 - Availability of Medical and Recreational Marijuana.Pdf. http://downloads.hindawi.com/journals/jad/2016/7193740.pdf. Accessed 7 Dec. 2019.

Shi, Yuyan, Kristin Meseck, et al. "Availability of Medical and Recreational Marijuana Stores and Neighborhood Characteristics in Colorado." *Journal of Addiction*, vol. 2016, 2016, pp. 1–7. *DOI.org (Crossref)*, doi:10.1155/2016/7193740.

Shi, Yuyan, Sharon E. Cummins, et al. "Medical Marijuana Availability, Price, and Product Variety, and Adolescents' Marijuana Use." *Journal of Adolescent Health*, vol. 63, no. 1, July 2018, pp. 88–93. *ScienceDirect*, doi:10.1016/j.jadohealth.2018.01.008.

Stephen E. Lankenau, Loni Philip Tabb, Avat Kioumarsi, Janna Ataiants, Ellen Iverson, Carolyn F. Wong. (2019) Density of Medical Marijuana Dispensaries and Current Marijuana Use among Young Adult Marijuana Users in Los Angeles. *Substance Use & Misuse* 54:11, pages 1862-1874.

Subica, Andrew M., et al. "The Geography of Crime and Violence Surrounding Tobacco Shops, Medical Marijuana Dispensaries, and off-Sale Alcohol Outlets in a Large, Urban

Low-Income Community of Color." *Preventive Medicine*, vol. 108, Mar. 2018, pp. 8–16. *ScienceDirect*, doi:10.1016/j.ypmed.2017.12.020.

Taiyun Wei and Viliam Simko (2017). R package "corrplot": Visualization of a Correlation Matrix (Version 0.84).

Available from https://github.com/taiyun/corrplot

Thomas, Crystal, and Bridget Freisthler. "Examining the Locations of Medical Marijuana Dispensaries in Los Angeles." *Drug and Alcohol Review*, vol. 35, no. 3, 2016, pp. 334–37. *Wiley Online Library*, doi:10.1111/dar.12325.

Thomas J. Leeper (2018). tabulizer: Bindings for Tabula PDF Table Extractor Library. R package version 0.2.2.

Tyndall, Justin. "Getting High and Low Prices: Marijuana Dispensaries and Home Values." *Real Estate Economics*, Oct. 2019, pp. 1540-6229.12302. *DOI.org (Crossref)*, doi:10.1111/1540-6229.12302.

United States, Congress, "American Community Survey: Data Suppression." *American Community Survey: Data Suppression*, 2016, pp. 1–12. pdf

U.S. Census Bureau; American Community Survey, 2015 American Community Survey 5-Year Estimates, generated by Sloan Stinnett; using the censusAPI; (3 December 2019)

Zeileis A (2004). "Econometric Computing with HC and HAC Covariance Matrix Estimators." _Journal of Statistical Software_, *11*(10), 1-17. doi: 10.18637/jss.v011.i10 (URL: https://doi.org/10.18637/jss.v011.i10).

Zhang, Kay, et al. "A Review of the Impact of Marijuana's Legalization on Colorado's Industrial Warehouse Lease Rates: How High Is High?" *Journal of Real Estate Literature*, vol. 25, no. 1, Jan. 2017, pp. 1–29. *aresjournals.org (Atypon)*, doi:10.5555/0927-7544.25.1.1.

Figures

Figure 1

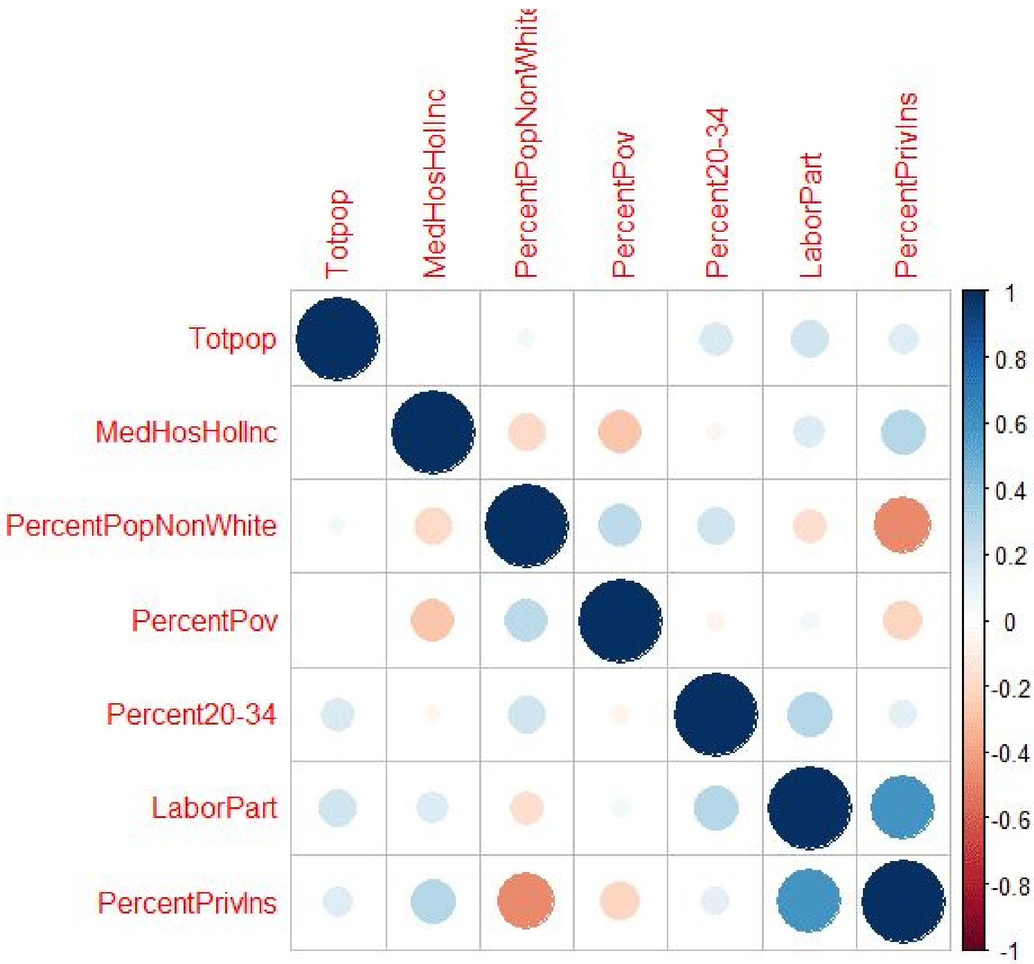


Table 1: Figure 2

	Dependent variable:	
	Freq	
	OLS	*coefficient test*
	(1)	(2)
Totpop	0.001***	0.001***
	(0.00001)	(0.0001)
MedHosHolInc	0.00001	0.00001
	(0.00003)	(0.00001)
PercentPopNonWhite	−0.023	−0.023
	(0.046)	(0.030)
PercentPov	0.019	0.019
	(0.118)	(0.077)
‘Percent20-34‘	−0.087	−0.087
	(0.098)	(0.110)
PercentPrivIns	−0.050	−0.050
	(0.052)	(0.043)
LaborPart	−0.084	−0.084
	(0.072)	(0.053)
Constant	9.572**	9.572*
	(3.903)	(5.127)
Observations	200	
R^2	0.977	
Adjusted R^2	0.976	
Residual Std. Error	6.355 (df = 192)	
F Statistic	1,149.510*** (df = 7; 192)	
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 1: Figure 3.

	Dependent variable:	
	Freq	
	OLS	*coefficient test*
	(1)	(2)
Totpop	0.001***	0.001***
	(0.00002)	(0.0001)
MedHosHolInc	−0.00000	−0.00000
	(0.00002)	(0.00001)
PercentPopNonWhite	−0.002	−0.002
	(0.037)	(0.022)
PercentPov	−0.078	−0.078*
	(0.095)	(0.044)
‘Percent20-34‘	0.059	0.059
	(0.080)	(0.099)
PercentPrivIns	0.005	0.005
	(0.042)	(0.035)
LaborPart	−0.041	−0.041
	(0.058)	(0.039)
I(Totpop *Totpop)	0.000***	0.000
	(0.000)	(0.000)
Constant	2.390	2.390
	(3.210)	(2.740)
Observations	200	
R^2	0.985	
Adjusted R^2	0.984	
Residual Std. Error	5.103 (df = 191)	
F Statistic	1,572.993*** (df = 8; 191)	

Table 1: Figure 4.

	Dependent variable:	
	Freq/(Totpop/100)	
	OLS	*coefficient test*
	(1)	(2)
MedHosHolInc	0.00000	0.00000
	(0.00000)	(0.00000)
PercentPopNonWhite	−0.001	−0.001
	(0.001)	(0.001)
PercentPov	−0.006***	−0.006***
	(0.002)	(0.002)
‘Percent20-34‘	−0.002	−0.002
	(0.002)	(0.002)
PercentPrivIns	−0.004***	−0.004**
	(0.001)	(0.002)
LaborPart	0.001	0.001
	(0.001)	(0.001)
Constant	0.314***	0.314***
	(0.074)	(0.109)
Observations	200	
R^2	0.107	
Adjusted R^2	0.079	
Residual Std. Error	0.123 (df = 193)	
F Statistic	3.862*** (df = 6; 193)	
Note:		*p<0.1; **p<0.05; ***p<0.01

Figure 5

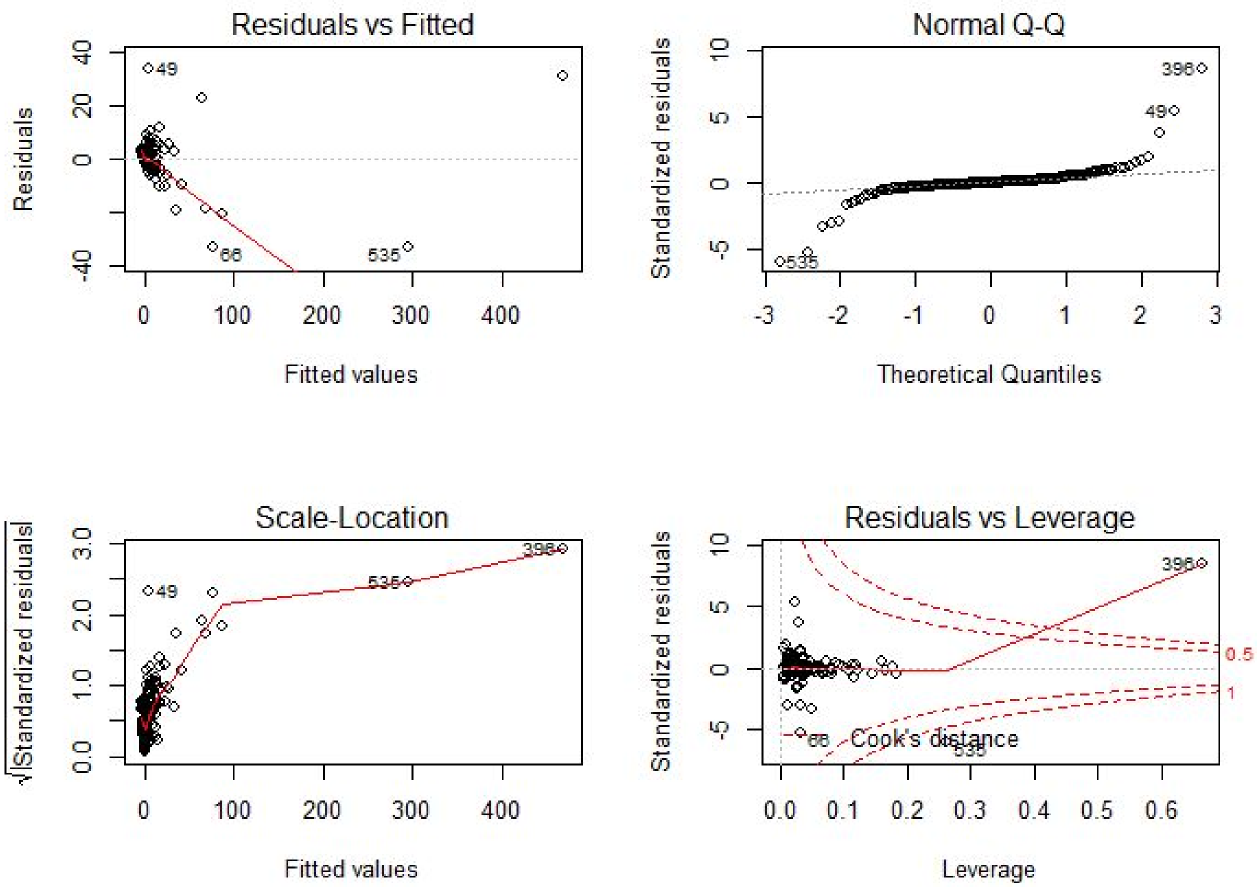

Residuals vs Fitted
Residuals
Fitted values
49
66
535
Normal Q-Q
Standardized residuals
Theoretical Quantiles
396
49
535
Scale-Location
√|Standardized residuals|
Fitted values
396
49
535
Residuals vs Leverage
Standardized residuals
Leverage
396
Cook's distance
0.5
1

Figure 6

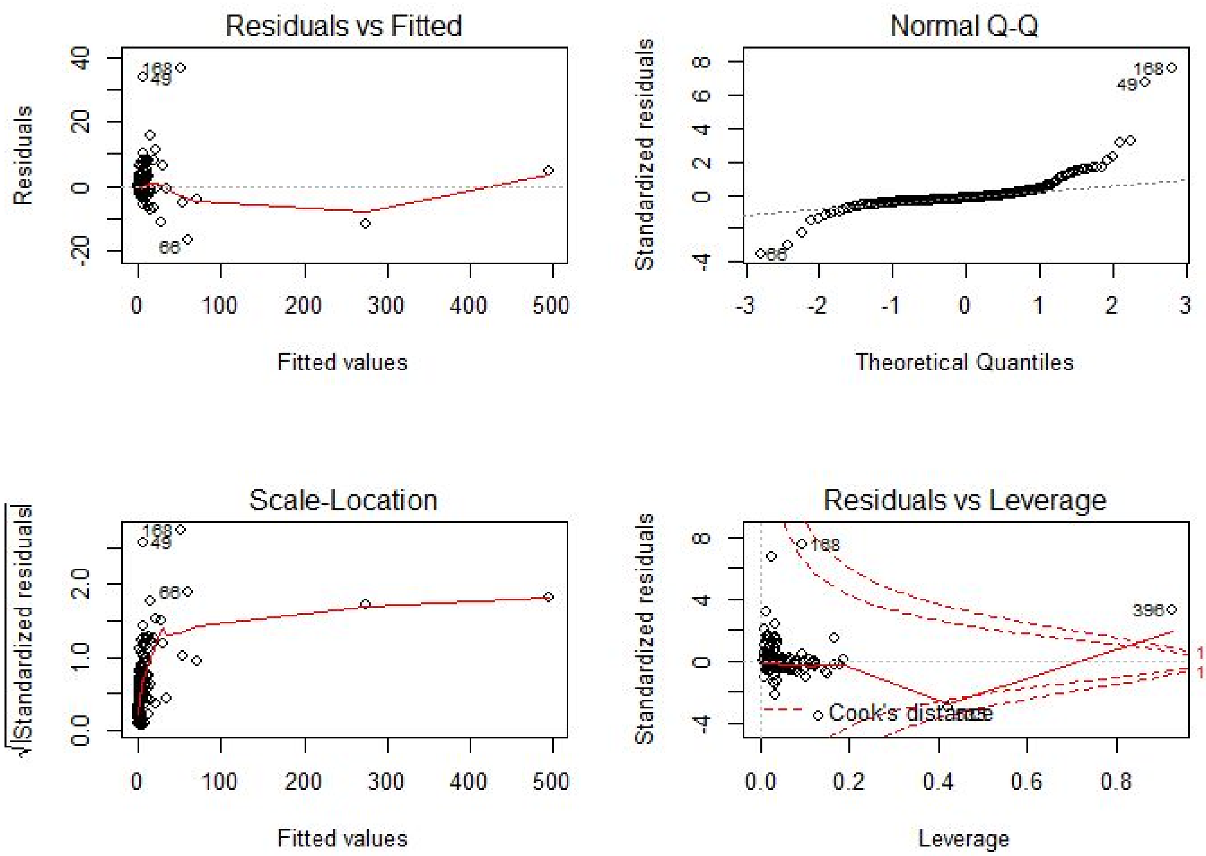
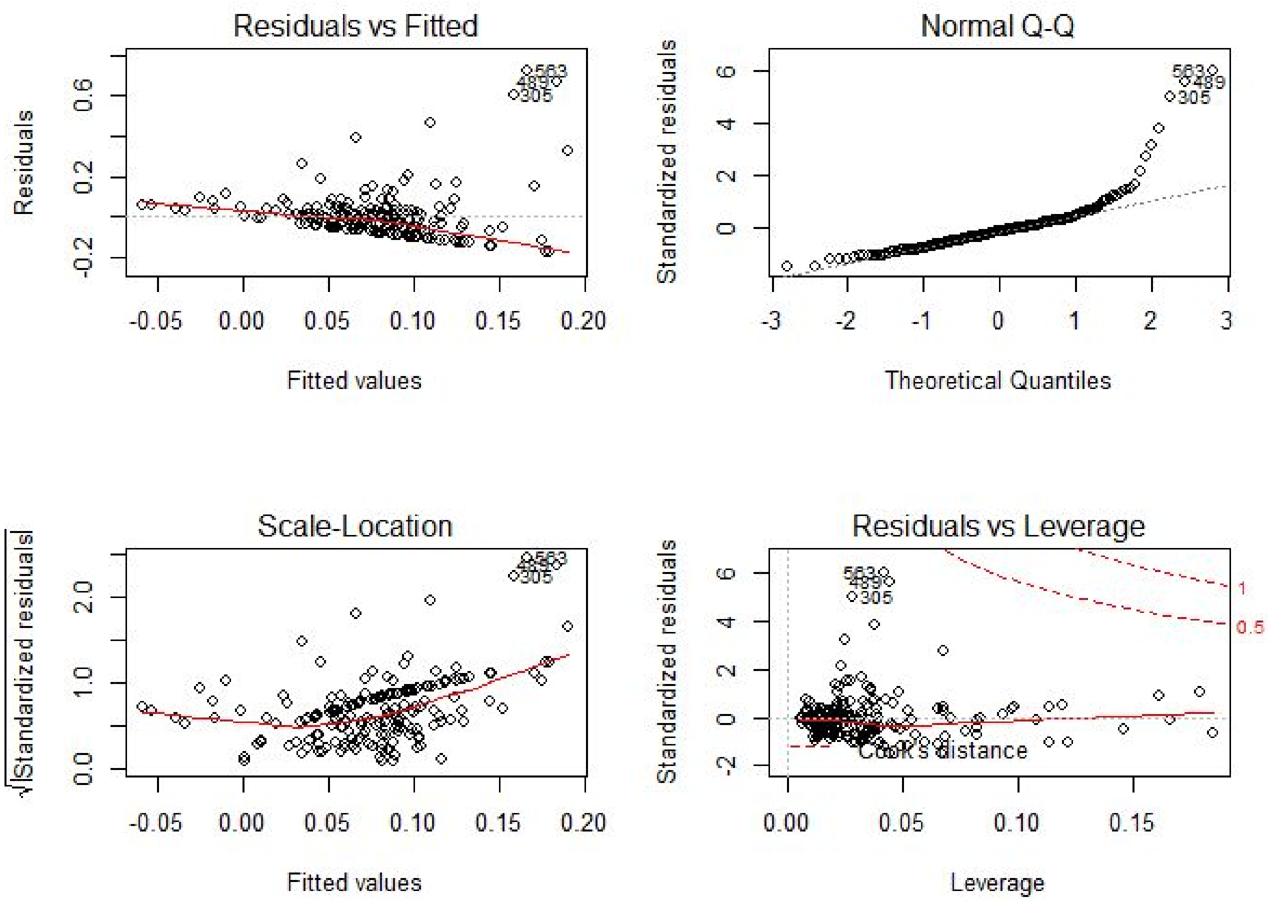

Residuals vs Fitted
Residuals
Fitted values
Normal Q-Q
Standardized residuals
Theoretical Quantiles
Scale-Location
√|Standardized residuals|
Fitted values
Residuals vs Leverage
Standardized residuals
Cook's distance
Leverage

Figure 7


Residuals vs Fitted
Residuals
Fitted values
563
489
305
Normal Q-Q
Standardized residuals
Theoretical Quantiles
Scale-Location
√|Standardized residuals|
Residuals vs Leverage
Cook's distance
Leverage
1
0.5